

Article

Beyond Supervised: The Rise of Self-Supervised Learning in Autonomous Systems

Hamed Taherdoost ^{1,2} 
¹ Department of Arts, Communications and Social Sciences, University Canada West, Vancouver, BC V6B 1V9, Canada; hamed.taherdoost@gmail.com or hamed@hamta.org

² GUS Institute, Global University Systems, London EC1N 2LX, UK

Abstract: Supervised learning has been the cornerstone of many successful medical imaging applications. However, its reliance on large labeled datasets poses significant challenges, especially in the medical domain, where data annotation is time-consuming and expensive. In response, self-supervised learning (SSL) has emerged as a promising alternative, leveraging unlabeled data to learn meaningful representations without explicit supervision. This paper provides a detailed overview of supervised learning and its limitations in medical imaging, underscoring the need for more efficient and scalable approaches. The study emphasizes the importance of the area under the curve (AUC) as a key evaluation metric in assessing SSL performance. The AUC offers a comprehensive measure of model performance across different operating points, which is crucial in medical applications, where false positives and negatives have significant consequences. Evaluating SSL methods based on the AUC allows for robust comparisons and ensures that models generalize well to real-world scenarios. This paper reviews recent advances in SSL for medical imaging, demonstrating their potential to revolutionize the field by mitigating challenges associated with supervised learning. Key results show that SSL techniques, by leveraging unlabeled data and optimizing performance metrics like the AUC, can significantly improve the diagnostic accuracy, scalability, and efficiency in medical image analysis. The findings highlight SSL's capability to reduce the dependency on labeled datasets and present a path forward for more scalable and effective medical imaging solutions.



Citation: Taherdoost, H. Beyond Supervised: The Rise of Self-Supervised Learning in Autonomous Systems. *Information* **2024**, *15*, 491. <https://doi.org/10.3390/info15080491>

Academic Editors: Jiaping Gui and Futai Zou

Received: 14 July 2024

Revised: 7 August 2024

Accepted: 15 August 2024

Published: 16 August 2024



Copyright: © 2024 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: self-supervised learning; medical imaging; area under the curve; image analysis; anomaly detection; classification; feature extraction; pre-training

1. Introduction

Most often, classification and regression challenges are tackled using supervised learning. The purpose of regression is to forecast a continuous numerical result, whereas the objective of classification is to predict a discrete class label. Linear and logistic regression, decision trees, support vector machines, neural networks, and linear regression are common supervised learning techniques [1–3].

It could be costly and time-consuming to gather and annotate such datasets. Models can underperform when presented with novel, unseen data if the training data need to be improved or are of low quality [4]. Supervised learning's success depends on feature engineering, which is challenging, calls for domain knowledge, and entails choosing and altering the most informative features [5]. Another typical problem with supervised learning is overfitting, which occurs when a model performs well on training data but needs to be improved on new data [6]. The predictions made by supervised learning algorithms could be biased if there are biases in the training data [7,8]. Research in machine learning is actively aimed at addressing algorithmic bias [9,10].

In recent years, self-supervised learning (SSL) has gained popularity as a potential solution to supervised learning problems, such as the enormous quantity of labeled data

required. SSL uses large amounts of available unlabeled data to develop helpful representations, unlike supervised learning, which relies on the human annotation of training data [11,12]. The basic tenet of SSL is to eliminate the requirement for human-provided labels by defining pretext tasks that can be addressed using the data's inherent structure and patterns [12–14]. Disease diagnosis and picture segmentation are examples of medical imaging tasks where SSL has outperformed strictly supervised methods [15].

The design of efficient pretext tasks and the possibility that the acquired representations may not be ideal for particular downstream tasks are two of the obstacles that SSL must overcome [16]. Researchers are investigating several approaches to combat these issues, such as merging SSL with other methods, including transfer learning, and creating more complex pretext tasks [13,16].

Common metrics include accuracy, precision, recall, and the F1 score, which provide distinct insights into model performance [17]. Accuracy evaluates a model's correctness, but precision and recall are more useful in imbalanced datasets by focusing on class performance [18]. The area under the curve (AUC) of the receiver operating characteristic (ROC) curve is useful in comparing models and handling imbalanced data because it summarizes the trade-off between the true positive rates (TPRs) and false positive rates (FPRs) across various thresholds [19]. These indicators help to select and optimize models and guarantee that they fulfill performance standards for key applications [20].

The AUC is the likelihood that a randomly picked positive instance would have a higher predicted score than a negative instance. The AUC runs from 0.5 to 1.0, with 0.5 suggesting a test no better than random chance and 1.0 indicating a flawless test that can discriminate between positive and negative instances. AUCs of 0.7 or higher are acceptable, 0.8 or higher are good, and 0.9 or higher are excellent [21–24]. The AUC provides a single scalar value that summarizes a binary classifier's performance, making it useful for model comparison and selection [24]. The AUC alone should not be used to evaluate models since it ignores decision thresholds, class imbalance, and fairness [25]. Diagnostic tests and classification models with a larger AUC are deemed more discriminative [26].

Despite SSL's rapid developments and transformational potential across domains, previous research generally lacks a thorough evaluation framework that robustly quantifies model performance [27,28]. Although some studies have investigated the learning of multiple pretext tasks for self-supervised feature learning, additional research is necessary to realize the potential of SSL completely [12]. Further research is required to address the challenges of diverse environments and scenarios when deploying SSL models in real-world autonomous systems [29].

While the AUC is a well-established metric in supervised learning, its use in SSL is underexplored [30]. This research systematically examines the inclusion of the AUC in SSL model assessment, providing a comprehensive understanding of their performance across applications. The significance of the AUC statistic in assessing SSL-based classification models is emphasized in this paper. It examines SSL applications in various domains, including image recognition and natural language processing, showcasing developments and a decreased dependency on labeled data. The article presents improved AUC-optimizing SSL frameworks and illustrates their effective applications in medical imaging. It also tackles the difficulty of acquiring high-quality tagged data and the requirement for more comprehensible models.

This paper will explore the idea and methods of SSL, explain the significance of the AUC metric in assessing SSL models, investigate the transformative applications of SSL in various domains, and present the methodology used for the critical review. Finally, it will discuss challenges and future directions and conclude on the significance of SSL and the function of the AUC in assessing its performance.

2. Concept and Techniques of Self-Supervised Learning

Key practicalities and workable profitabilities associated with functional information prediction have emerged with the advent of the SSL concept. These have contributed to

desirable outcomes in satisfying precision, computational effort flexibility, time efficiency, and cost-effectiveness [31], especially for diverse disciplines without prior annotated documented databases. After optimizing the surrogate objective, the pre-trained model can be used as a feature extractor to feed downstream supervised tasks [32,33]. SSL approaches have demonstrated promising outcomes in practice, but their theoretical foundations are unknown [34].

Figure 1 shows how the SSL framework preprocesses raw data for analysis. After creating pseudo-labels and pretext tasks, a self-supervised task is produced. This challenge starts with model training using contrastive learning, generative models, and predictive models. After training, the model learns features through extraction and representation. Later, tasks like image recognition, NLP, robotics, and autonomous vehicles use these learned properties. The AUC, accuracy, precision, and recall are used to evaluate the model's performance in various tasks.

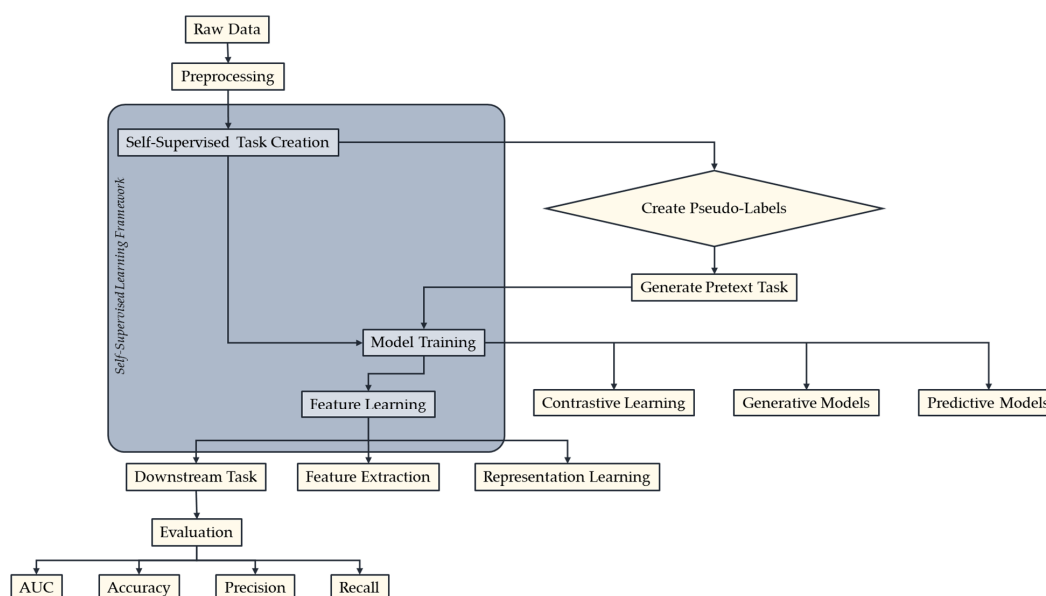


Figure 1. SSL concept.

Contrastive, generating, and predictive tasks are the three primary types of SSL pretext tasks (Table 1) [35]. These three types of SSL have helped computer vision and natural language processing models to develop powerful representations from unlabeled data. The learned representations can be tailored to downstream activities.

Table 1. SSL pretext tasks.

Category	Description	Examples	Sources
Contrastive Learning	Methods that train models to distinguish between related and unrelated data samples.	SimCLR: A simple framework for the acquisition of contrastive visual representations MoCo: Contrastive momentum for unsupervised learning of visual representations	[11,27,28,36]
Generative Models	Methods that train models to generate or reconstruct the input data, capturing the underlying data distribution.	Variational autoencoders (VAEs): Generative models that learn a latent representation of the data Generative adversarial networks (GANs): Adversarial training of a generator to produce realistic samples	[35–38]

Table 1. Cont.

Category	Description	Examples	Sources
Predictive Models	Tasks that involve predicting some aspect of the input data require the model to understand and capture relevant features.	Predicting relative positions of image patches	[35]
		Solving jigsaw puzzles formed from image patches	
		Predicting rotations applied to images	

3. Transformative Applications of Self-Supervised Learning

The constraints of traditional machine learning models have been addressed by SSL, a promising deep learning technique that trains neural networks using enormous volumes of unlabeled data [39]. SSL techniques address challenges such as data sparsity and the necessity of extensive human effort in annotating datasets by utilizing immense volumes of unlabeled data to train neural networks [15,40,41]. These reviews underscore the potential of SSL to improve the accuracy of prediction, maintenance prognostics, and illness diagnosis, all while reducing the computational budgets, time, and storage requirements. The advancement of SSL methodologies, such as contrastive learning, generative learning, and adversarial learning, has allowed models to learn complex patterns from unlabeled data, resulting in enhanced performance in tasks such as semantic segmentation, object detection, and image classification [40,41]. The transformative impact of SSL on a variety of disciplines is emphasized by the synthesis of recent breakthroughs, which opens the door for additional innovations in deep learning [42].

3.1. Image Recognition

Clinical decision-making could be enhanced by machine learning and deep learning approaches, which could ultimately result in better patient outcomes [43]. SSL holds significant potential for revolutionary image recognition applications. SSL algorithms like self-supervised deep learning within self-distillation with no labels (DINO) [44] have achieved state-of-the-art medical image classification performance with 1–10% of the labeled training data needed by fully supervised methods. Medical imaging is especially affected, where labeled data are rare and expensive. On downstream tasks, self-supervised pre-training outperforms fully supervised pre-training for object detectors [45]. The learned representations are more reliable and generalizable. SSL systems like PixPro [15] learn sophisticated visual representations from unlabeled data to promote semantic segmentation. This helps to resolve dense pixel-level annotation issues. Azizi et al. [46] developed a representation learning technique for medical imaging machine learning models that reduces ‘out of distribution’ performance issues and enhances the model robustness and training efficiency. REMEDIS (‘Robust and Efficient Medical Imaging with Self-Supervision’) uses large-scale supervised transfer learning on natural pictures and intermediate contrastive SSL on medical images with minimum task-specific customization. REMEDIS enhanced the in-distribution diagnosis accuracies by 11.5% compared to strong supervised baseline models and required only 1–33% of the data for retraining to match supervised models retrained using all available data.

3.2. Natural Language Processing (NLP)

Bidirectional Encoder Representations from Transformers (BERT) and T5 models have revolutionized NLP through SSL. These models develop rich contextual representations of language by being pre-trained on huge unlabeled text corpora, employing self-supervised aims like masked language modeling. The learned representations can be fine-tuned on downstream NLP tasks, including text categorization, question answering, and language production, improving the performance significantly [12]. BERT and T5’s achievement has spurred the study of the transformation of self-supervised objectives. Alternatives to masked language modeling include random token substitution, cluster-based random

token substitution, and swapped language modeling. The structured alignment of pre-training objectives with downstream tasks may help to reduce the labeled data needs [47].

Zhou et al. [48] improved text categorization across 17 datasets using self-supervised regularization. On BERT's self-supervised pre-training, Gururangan et al. [49] suggested task-adaptive and domain-adaptive pre-training to specialize the model for target tasks and domains. Sun et al.'s [50] ERNIE 2.0 system trains representations on self-supervised tasks such as masked language modeling, named entity identification, and sentence sorting.

3.3. Robotics

STERLING, a unique method for the learning of terrain representations from unconstrained robot experience, enables robust off-road navigation without labeled data or expert demonstrations [51]. Regarding the self-supervised prediction of human interaction intent, researchers have developed learning-based methods to estimate the probability that a person will interact with a service robot before the encounter occurs by learning relevant representations from sensor data [52]. Using self-supervised video learning, time-contrastive networks learn visual representations for robotic system control without manual labeling. Deep learning and large-scale data collection have helped robots to learn robust grasping skills in a self-supervised manner [53].

3.4. Autonomous Vehicles

SSL approaches have shown promise in autonomous vehicle tasks such as long-range traversable area segmentation, moving obstacle instance segmentation, long-term moving obstacle tracking, and depth map prediction [29,54,55]. SSL approaches allow models to learn in real time and adapt to changing surroundings without human annotation, offering an alternative to supervised learning. SSL in autonomous vehicles may improve perception, enabling them to traverse complicated settings more accurately and efficiently. SSL approaches can help autonomous vehicles to construct strong perception systems that can tolerate unexpected sensor data fluctuations, making autonomous driving safer and more dependable. Bojarski's [56] end-to-end learning for self-driving automobiles showed that SSL improves autonomous systems. SSL can help deep learning-based autonomous driving algorithms to become more advanced and adaptive [57].

4. Importance of Area under the Curve in Machine Learning

Binary classification models are often evaluated using the area under the ROC curve (AUC). It reflects the likelihood that a randomly picked positive example will be ranked higher than a negative example [58–60]. The AUC is calculated by plotting the TPR versus the FPR at various classification levels (Figure 2). The ratio of accurately anticipated positive cases to the total positive instances is the recall or TPR. The FPR is the ratio of miscalculated positives to negatives.

Recall, sometimes called the TPR, is the percentage of true positive cases that the model properly identifies. It is computed as follows:

$$TPR = \frac{TP}{TP + FN} \quad (1)$$

The actual positive occurrences mistakenly projected as negative are false negatives (FN), and the correctly predicted positive examples are true positives (TP).

Conversely, the FPR denotes the percentage of true negative cases mistakenly categorized as positive. It is computed as follows:

$$FPR = \frac{FP}{FP + TN} \quad (2)$$

True negatives (TN) are the accurately predicted negative cases, and false positives (FP) are the real negative instances mistakenly projected as positive.

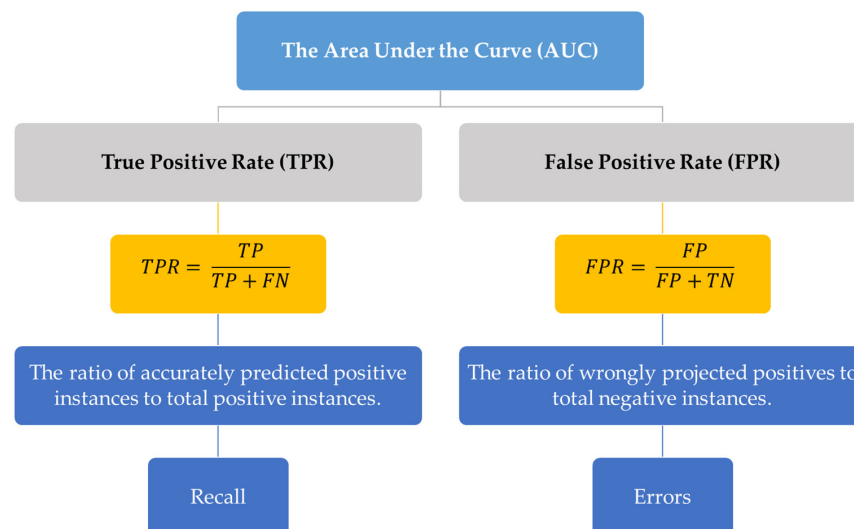


Figure 2. AUC through TPR and FPR.

The AUC evaluates the model's performance across all feasible thresholds; it is not dependent on any particular threshold. This attribute is especially advantageous for SSL models, in which the ideal threshold for classification might need to be more readily discernible. The AUC, which offers a comprehensive performance evaluation, facilitates the comprehension of the model's overall discriminatory capabilities independently of any particular threshold [61–63].

The interpretability of the AUC is highly regarded as it offers a concise summary of a model's overall capability to differentiate between positive and negative classes in the form of a single scalar value. The AUC is a fundamental machine learning statistic that summarizes a model's ability to differentiate between positive and negative classes using a scalar number from 0 to 1. As a result of this interpretation's simplicity, the AUC is an intuitive metric that facilitates the comparison of the performance of various models. Furthermore, the AUC offers a comprehensive assessment of model performance by incorporating both the TPR and FPR. This enables it to account for both classification errors, which is especially valuable when the expenses associated with false positives and false negatives differ [21,22,24,64–66].

As a metric for the assessment of the performance of classification models, including those trained using SSL techniques, the area under the ROC curve, or AUC, is extensively employed [67–69]. The AUC is a metric that offers a thorough evaluation of a model's capability to differentiate between positive and negative classes, taking into consideration the compromise between the FPR (specificity) and the TPR (sensitivity) [64,70–72].

The AUC holds significant relevance within the domain of SSL due to its capacity to quantify the model's aptitude for the acquisition of meaningful representations from unlabeled data, a fundamental goal of SSL methodologies [68]. Increasing the AUC during SSL training can motivate the model to acquire more discriminatory features among various classes, enhancing its classification task performance [67,69].

Several studies, including SSLROC1 and SSLROC2, have introduced SSL frameworks that optimize the AUC directly. These frameworks have demonstrated superior performance to alternative supervised and semi-supervised AUC optimization approaches [69,73]. The AUC-CL approach incorporates a batch-size-robust AUC maximization goal for SSL, which exhibits enhanced efficacy compared to conventional contrastive learning methods such as SimCLR, particularly when the batch size is reduced [68].

5. Methodology

This research uses a systematic review to examine how SSL transforms diverse domains and how the AUC is important in SSL model evaluation. An organized and compre-

hensive literature review (May 2024) focuses on SSL studies using the AUC as an evaluation tool.

Keyword Query:

("self-supervised learning" OR "self-supervised model" OR "self-supervision" OR "SSL") AND ("supervised learning" OR "unsupervised learning" OR "supervised model" OR "unsupervised model") AND ("Area Under the Curve" OR "AUC").

A total of 164 preliminary results from ScienceDirect (32 results), IEEE Xplore (129 results), and the ACM Digital Library (3 results) were obtained from the search. After applying the inclusion and exclusion criteria (Table 2), 51 papers were included for assessment. After a comprehensive review of the abstracts and titles, 34 papers were found to have possible relevance to the research. The methodical process from the first database search to the identification of possibly eligible articles for review inclusion is depicted in Figure 3.

Table 2. Selection criteria for research papers.

Selection Criteria	Details
Inclusion Criteria	<ul style="list-style-type: none"> - Journal papers - Studies focusing on SSL techniques in image recognition, NLP, robotics, and autonomous vehicles - Papers that utilize AUC as performance metric for SSL models and mention exact number of AUCs
Exclusion Criteria	<ul style="list-style-type: none"> - Studies not published between 2019 and 2024 - Non-peer-reviewed articles - Studies not involving SSL or AUC - Articles that do not provide sufficient methodological details

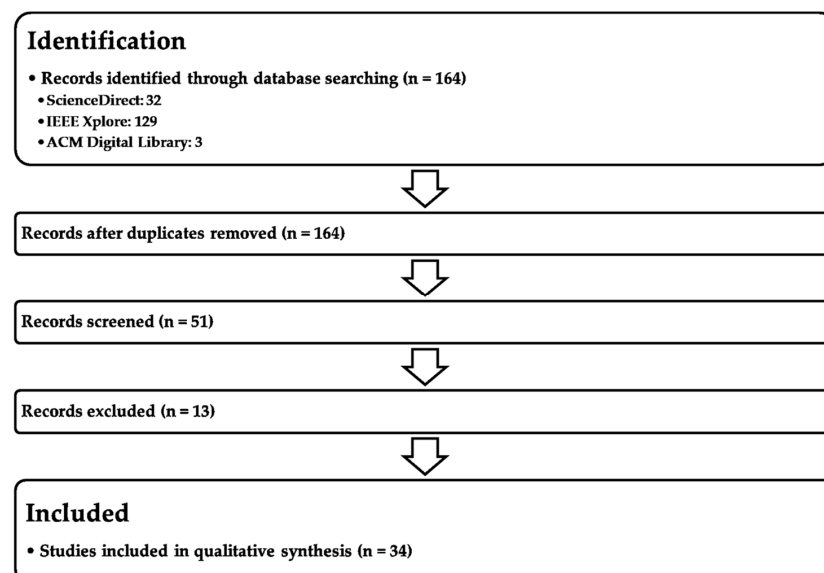


Figure 3. Process of selecting papers for review.

The AUC provides a comprehensive evaluation of model performance by taking into account the trade-offs between specificity and sensitivity. Nevertheless, our research emphasizes the AUC; however, a more comprehensive evaluation framework can be provided by other metrics, such as the precision, recall, and F1 score.

6. Critical Review of Self-Supervised Learning Models Using AUC

Several medical imaging investigations have shown that general-purpose SSL techniques are effective and versatile. Dong et al. [74] suggested a method to convert a convolutional neural network (CNN) designed for registered pictures to work with unregistered images, demonstrating competitive performance in large vessel occlusion (LVO) detection using computed tomographic angiography (CTA) data. The electrocardiogram-masked autoencoder (ECG-MAE), a unique generative self-supervised pre-training method for the learning of spatiotemporal representations from multi-lead ECG signals, yields improved multi-label classification performance, according to Hu et al. [75]. Zhao et al.'s [76] AddNet-Supervised Contrastive Learning (SCL) model reduced the computing costs by using addition instead of large multiplication in the convolution process, obtaining competitive outcomes in clinical practice, particularly in epilepsy diagnosis. Lu and Dai [77] introduced a CT-based COVID-19 recognition system that applied two-phase contrastive SSL to the backbone network to classify multiple labels with high accuracy and a small amount of training data.

Sun et al. [78] developed the TSRNet model using transfer learning with self-supervised (TS) pre-training to create a CNN based on an attention mechanism and deep residual network (RANet). It outperformed other models in classifying the lung CT images of suspected COVID-19 patients. Pascual et al. [79] used SSL in wireless endoscopic recordings to increase the polyp detection rates by inferring the inherent structure. General-purpose SSL techniques can improve medical image analysis for various diagnostic tasks, from cardiovascular anomaly detection to infectious disease identification, enabling more efficient and robust artificial intelligence (AI)-powered healthcare systems.

Figure 4 classifies the works by methodology: self-supervised, semi-supervised, weakly supervised, transfer, generative, and hybrid. It also describes their uses in medical imaging (lung CT, whole-slide image (WSI) analysis, positron emission tomography (PET)/CT, knee magnetic resonance (MR) films, and cardiac ECG), defect identification, pathology, classification (COVID-19 and subcentimeter solid pulmonary nodule (SSPN) malignancy), and segmentation.

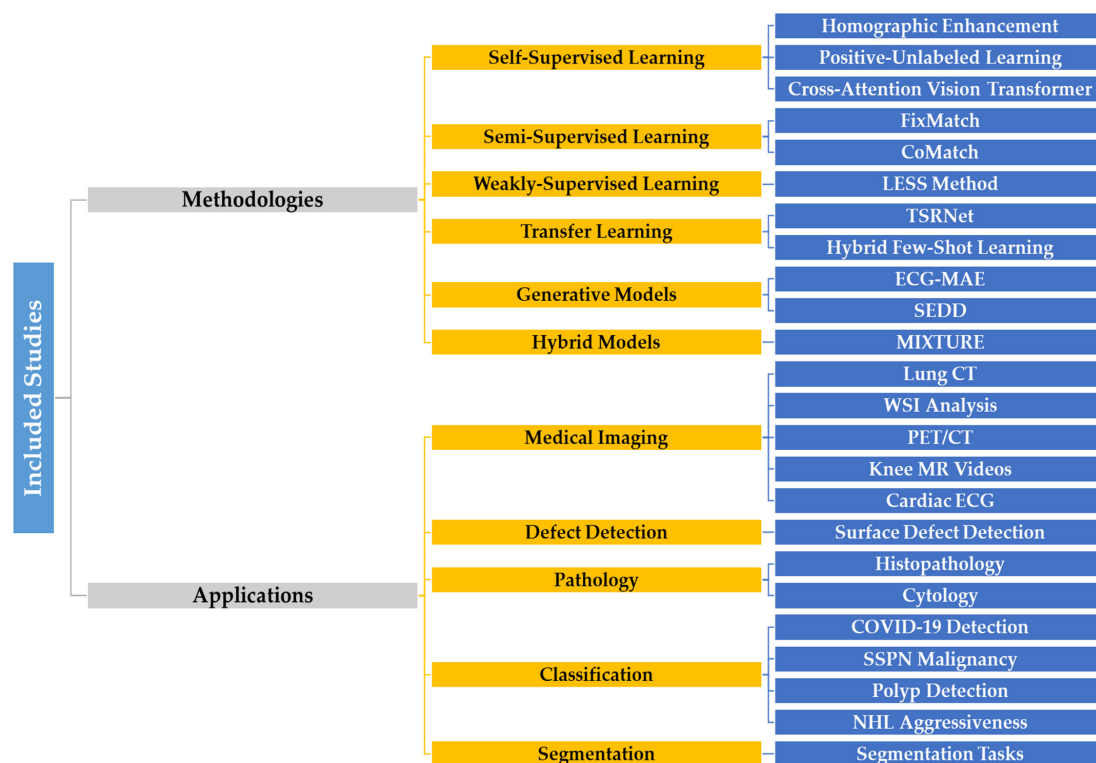


Figure 4. Overview of methodologies and applications in included studies.

6.1. Diagnostic Imaging Classification

Numerous studies have shown that deep learning systems can distinguish medical disorders. Wongchaisuwat et al. [80] and Liu et al. [81] used optical coherence tomography (OCT) images to apply SSL to distinguish polypoidal choroidal vasculopathy (PCV) and wet age-related macular degeneration (AMD). Wongchaisuwat et al. [80] achieved an AUC of 0.71, whereas Liu et al. [81] employed CT scans to distinguish malignant and benign SSPNs with outstanding diagnostic accuracy. Using PET/CT images, Xu et al. [82] proposed a hybrid few-shot multiple-instance learning model to predict non-Hodgkin's lymphoma (NHL) aggressiveness with high accuracy and AUC ratings. Sun et al. [78] used lung CT images to diagnose COVID-19 using transfer learning with self-supervised pre-training, exceeding previous accuracy, recall, and AUC models. Perumal and Srinivas' [83] DenSplitnet model for COVID-19 identification in chest CT scans used dense blocks and SSL and performed well across many datasets. Manna et al. [84] suggested an SSL system for video-based injury categorization in MR imaging, achieving high accuracy and AUC scores, demonstrating the potential of self-supervised methods to improve diagnostic imaging.

6.2. Defect Detection and Segmentation

Recent research has shown that new defect identification and segmentation approaches have improved the picture anomaly detection accuracy and efficiency. The Self-Supervised Efficient Defect Detector (SEDD) by Xu et al. [85] uses SSL and image segmentation to detect defects competitively without annotated data. They use homographic improvement and lightweight structures with attention modules to achieve great accuracy and a low computational overhead. The Growth Threshold for Pseudo-Labeling (GTPL) and Pseudo-Label Dropout (PLD) techniques improve semi-supervised classification models like FixMatch and CoMatch in detecting and segmenting medical picture abnormalities, according to Zhou et al. [86,87]. Their strategy improves the pseudo-label quality and segmentation accuracy by dynamically modifying the thresholds and smoothing the labels.

6.3. Self-Supervised Learning in Pathology

In pathology, SSL algorithms have become strong tools that provide novel ways to extract pathologically relevant data with little annotation work. To create deep learning models, Uegami et al. [88] presented MIXTURE (Human-in-the-Loop Explainable AI through the Use of Recurrent Training), a technique that combines expert pathologist input with SSL. Similarly, Wongchaisuwat et al. [80] used SSL techniques to create an automated classification model that used macula optical coherence tomography (OCT) images to distinguish between PCV and wet AMD. Zhao et al. [89] proposed a weakly supervised label-efficient WSI screening method (LESS) for cytological WSI analysis. This method uses a Cross-Attention Vision Transformer (CrossViT) and variational positive unlabeled (VPU) learning to classify WSI accurately using only slide-level labels. To increase the labeling consistency in the COUGHVID dataset for a variety of cough sound classification tasks, such as differentiating between COVID-19 and healthy coughs, Orlandic et al. [90] used a semi-supervised learning technique.

7. AUC Evaluation

SSL approaches excel in numerous tasks. Self-supervised masking (SSM) for anomaly detection, ViT-based models for cervical OCT image classification, and SEDD for surface defect detection excel in terms of their AUC scores and outperform state-of-the-art methods. Second, hybrid SSL methods offer significant improvements. This tendency has been shown using SSL with multiple instance learning (MIL) for WSI classification and TS for COVID-19 classification. These hybrid models generally outperform standalone SSL or standard supervised approaches, showing that various learning paradigms can improve the prediction abilities.

Certain activities are performed well using medical-specific approaches. Contrast-shifted instances via patch-based percentile (CSIP) for automatic lung shadowing iden-

tification and the mixed self- and weakly supervised learning framework for medical imaging abnormality detection have higher AUC scores than previous approaches. This shows that custom approaches to medical data can increase the diagnostic accuracy and anomaly identification, thereby improving the clinical results. Table 3 shows the AUC values and comparisons of self-supervised and semi-supervised learning approaches in medical imaging tasks.

Table 3. Performance comparison of self-supervised and semi-supervised learning methods in medical imaging tasks.

Method	Application	AUC	Comparison Results	Reference
General-purpose contrastive SSL	LVO detection in CTA data	0.88	Competitive performance compared to the teacher model	[74]
ECG-MAE: Generative self-supervised pre-training	Multi-label ECG classification	0.9474 (macro-averaged AUC)	Exceeded prior studies in downstream performance	[75]
AddNet-SCL	Clinical diagnosis using EEG signals	94.2%	Competitive results compared to state-of-the-art methods	[76]
COVID-19 recognition system based on CT images	COVID-19 diagnosis	0.989	Achieved high recognition accuracy with limited training data	[77]
TS	COVID-19 classification from lung CT images	1	Highest accuracy achieved compared to existing models	[78]
Pre-training method based on transfer learning with SSL	Polyp detection in wireless endoscopy videos	$95.00 \pm 2.09\%$	Achieved state-of-the-art results in polyp detection	[79]
SSL technique for OCT image classification	Differentiation between PCV and wet AMD from OCT images	0.71	Desirable performance with a small proportion of labeled data	[80]
Self-supervision pre-training-based fine-grained network	Differentiating malignant and benign SSPNs	0.964 (internal testing set), 0.945 (external test set)	Robust performance in predicting SSPN malignancy	[81]
Hybrid few-shot multiple-instance learning model with SSL	Diffuse large B-cell lymphoma (DLBCL) versus follicular lymphoma (FL) classification in PET/CT images	0.795 ± 0.009	Outperformed typical counterparts in NHL aggressiveness prediction	[82]
DenSplitnet: Dense blocks with SSL	COVID-19 classification from chest CT scans	0.95	Outperformed other methods in COVID-19 diagnosis	[83]
SSL approach for MR video classification	Classification of anterior cruciate ligament tear injury from knee MR videos	0.848	Achieved reliable and explainable performance	[84]
SEND	Surface defect detection	98.40% (average)	Achieved competitive performance with minimal computational consumption	[85]
GTPL and PLD	Semi-supervised skin lesion diagnosis	89.19–94.76%	Improved semi-supervised classification performance	[86]
FixMatch-LS and FixMatch-LS-v2	Medical image classification (skin lesion)	91.63–95.44%	Improved performance with label smoothing and consistency constraints	[87]

Table 3. Cont.

Method	Application	AUC	Comparison Results	Reference
MIXTURE: Human-in-the-loop explainable AI	Usual interstitial pneumonia diagnosis from pathology images	0.90 (validation set), 0.86 (test set)	Achieved high accuracy with an explainable AI approach	[88]
LESS	Cytological WSI analysis	96.86%	Outperformed state-of-the-art MIL methods on pathology WSIs	[89]
Semi-supervised learning for cough sound classification	COVID-19 versus healthy cough sound classification	0.797	Increased labeling consistency and improved classification performance	[90]
Morph-SSL	Longitudinal OCT scans for prediction of nAMD conversion	0.779	Outperformed end-to-end and pre-trained models	[91]
AntiViralDL	Predicting virus–drug associations	0.8450	Outperformed four benchmarked models	[92]
Multi-modal MRI based on contrastive learning	Diagnosis of major depressive disorder (MDD)	0.7309	Achieved 73.09% AUC	[93]
SSM	Anomaly detection and localization	0.983 (Retinal-OCT), 0.939 (MVTec AD)	Outperformed several state-of-the-art methods	[94]
Self-supervised ViT-based model	Cervical OCT image classification	0.9963 \pm 0.0069	Outperformed Transformers and CNNs	[95]
Masked relation learning	DeepFake detection	+2% AUC improvement	Outperformed state-of-the-art methods	[96]
Contrastive multi-view composite graph convolutional networks (CMV-CGCN)	Autism spectrum disorder (ASD) classification	0.7338	Outperformed state-of-the-art methods	[97]
Bayesian collaborative learning (BCL)	WSI classification	95.6% (CAMELYON16), 96.0% (TCGA-NSCLC), 97.5% (TCGA-RCC)	Outperformed all compared methods	[98]
Temporal cycle consistency (TCC)	Predicting pregnancy likelihood from developing embryo videos	0.64	Outperformed time alignment measurement (TAM)	[99]
Multiscale two-branch feature fusion	Self-supervised image anomaly detection	98.82%	Outperformed existing anomaly detection methods	[100]
Self-supervised transfer learning based on domain adaptation (SSTL-DA)	Benign–malignant lung nodule classification	95.84%	Achieved competitive classification performance	[101]
Coupling SSL and MIL for WSI classification	WSI classification	0.801 (Cohen’s kappa with 450 patch labels)	Achieved competitive performance with SSL and MIL baselines	[102]
CSIP	Automatic detection of diseased lung shadowing	0.96 (average AUC)	Improved diagnostic performance compared to existing methods	[103]
Mixed self- and weakly supervised learning framework	Abnormality detection in medical imaging	Up to 0.86 (image-wise AUC)	Competitive results versus multiple state-of-the-art methods	[104]
Pathology-specific self-supervised models	Classification of gigapixel pathology slides	0.80 (mean AUC)	Achieved competitive performance in immune subtype classification	[105]

Table 3. Cont.

Method	Application	AUC	Comparison Results	Reference
MIL-based framework with self-supervised pre-training	Cancer classification from whole-slide images	93.07% (accuracy), 95.31%	Outperformed existing methods	[106]
Automated differentiation between PCV and wet AMD	Optical coherence tomography (OCT) image analysis	0.71	Desirable performance compared to traditional supervised learning models	[107]

Techniques for SSL, such as SSM, have demonstrated impressive effectiveness in improving the performance on a range of benchmarks. SSM's ability to obtain high AUC scores in medical imaging anomaly detection tasks highlights its potential in accurately detecting abnormalities. Methods like TS and SSTL-DA demonstrate strong generalizability across different datasets and tasks. This adaptability is encouraging for real-world applications where the capacity to adjust to different data distributions is crucial. Even so, certain approaches have limitations due to task specificity, which restricts their cross-domain applicability. For example, these methods may work well in detecting medical anomalies but not for applications like drug–virus prediction. There are difficulties in gaining clinical acceptance and comprehension due to the uninterpretability of some sophisticated models, such as dense blocks with SSL or self-supervised ViT-based architectures. Moreover, the adoption of many state-of-the-art methods may be hindered by the significant data requirements and computational complexity associated with them, especially in healthcare settings with limited resources, where access to large datasets and high-performance computing resources is scarce.

8. Challenges and Future Directions

The utilization of artificial intelligence models has the potential to enhance diagnosis and treatment using early detection, which is a critical step in the enhancement of survival rates. Conversely, these models necessitate intricate information regarding patient severity and lesion characteristics in the form of lexical variables, which can be obtained through the semantic annotation of medical images [108]. High-quality tagged SSL data are difficult to obtain, especially in medical imaging, where annotations are expensive and time-consuming. Effective SSL requires data diversity and representativeness. Many SSL approaches, especially deep learning ones, need a large amount of processing power for training, making them inaccessible to researchers with limited computational resources. The SSL algorithms' efficiency must be improved without affecting their performance. Understanding SSL models' predictions and making their decisions clear and interpretable becomes more difficult as they become more complicated. AI judgments affect patient outcomes in medical imaging, making this crucial.

- The AUC, a widely used metric in SSL model evaluation, may only sometimes accurately reflect the performance, particularly in datasets with imbalances. Meaningful evaluation requires metrics like the AUC to appropriately assess model performance.
- Precise and consistent annotations are vital to ensure SSL models' accuracy in medical imaging datasets. Poorly annotated data can significantly skew model predictions, highlighting the importance of high-quality labeled data.
- Medical imaging datasets often lack labeled data, especially for rare diseases, hindering SSL model training. Addressing the data scarcity and imbalances is crucial for SSL techniques to learn minority class representations effectively.
- SSL approaches based on deep learning often feature complex architectures, necessitating significant computational resources for training and inference. Simplifying the model structures or adopting more efficient techniques can mitigate the computational complexity and improve the accessibility.

- Clinicians must trust and comprehend SSL models for widespread clinical use. Transparent model creation, validation, and interpretable predictions are essential in facilitating clinical adoption.

Researchers can use active learning or crowdsourcing annotation tasks to construct SSL models to improve the quality and quantity of labeled medical imaging data. Future research can integrate SSL approaches with robust evaluation measures like the AUC into clinical workflows to improve medical imaging diagnosis and therapy planning. More efficient and scalable SSL algorithms, especially for resource-constrained contexts, can help medical imaging applications to embrace SSL. Lightweight model architectures, optimization, and distributed computing are explored.

Integrating SSL approaches with multi-modal imaging data and using the AUC to evaluate multi-modal SSL models can improve diagnostic accuracy and patient care. SSL techniques for longitudinal and dynamic imaging analysis, such as tracking illness development or therapy responses, can reveal disease processes and individualized patient care strategies.

Before broad implementation, SSL models must undergo extensive clinical validation tests to determine their real-world performance and clinical utility in medical imaging. Future research should prioritize collaboration between computer scientists, clinicians, and medical experts to evaluate SSL models in varied clinical situations. Transferring knowledge from pre-trained SSL models to new imaging modalities or clinical domains using transfer learning and domain adaptation can speed up model development and increase generalization in real-world applications.

9. Conclusions

SSL methods consistently outperform baseline and supervised learning methods in medical imaging tasks. SSL models have been used for illness diagnosis, anomaly detection, medical picture categorization, and treatment prediction. Tasks like predicting virus–drug relationships, classifying pathology slides, detecting neurological illnesses, and distinguishing medical situations have been improved.

SSL models commonly have imbalanced datasets or uncertain class boundaries, making the AUC a good performance robustness indicator. SSL models' consistently excellent AUC values across medical imaging tasks demonstrate their ability to handle complicated medical data and capture key information for reliable predictions. SSL models outperform state-of-the-art methodologies and classic supervised learning methods, potentially improving medical image processing.

Cross-validation methods like splitting datasets into training and testing sets to validate the reported AUC scores and ensure the robustness and reliability of findings, as well as a sensitivity analysis to examine the impact of key parameters (e.g., training data size, model architecture) on the AUC scores and identify potential bias and variability, may be used in future research.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Tufail, S.; Riggs, H.; Tariq, M.; Sarwat, A.I. Advancements and Challenges in Machine Learning: A Comprehensive Review of Models, Libraries, Applications, and Algorithms. *Electronics* **2023**, *12*, 1789. [\[CrossRef\]](#)
2. Sarker, I.H. Machine learning: Algorithms, real-world applications and research directions. *SN Comput. Sci.* **2021**, *2*, 160. [\[CrossRef\]](#)
3. Taherdoost, H. Machine learning algorithms: Features and applications. In *Encyclopedia of Data Science and Machine Learning*; IGI Global: Hershey, PA, USA, 2023; pp. 938–960.
4. Fink, O.; Wang, Q.; Svensen, M.; Dersin, P.; Lee, W.-J.; Ducoffe, M. Potential, challenges and future directions for deep learning in prognostics and health management applications. *Eng. Appl. Artif. Intell.* **2020**, *92*, 103678. [\[CrossRef\]](#)
5. Liu, B. Supervised Learning. In *Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data*; Liu, B., Ed.; Springer: Berlin/Heidelberg, Germany, 2011; pp. 63–132.
6. Oliver, A.; Odena, A.; Raffel, C.; Cubuk, E.; Goodfellow, I. Realistic evaluation of semi-supervised learning algorithms. In Proceedings of the International Conference on Learning Representations, Vancouver, BC, Canada, 30 April–3 May 2018; pp. 1–15.
7. Ferrara, E. Fairness and bias in artificial intelligence: A brief survey of sources, impacts, and mitigation strategies. *Sci* **2023**, *6*, 3. [\[CrossRef\]](#)
8. Gianfrancesco, M.A.; Tamang, S.; Yazdany, J.; Schmajuk, G. Potential biases in machine learning algorithms using electronic health record data. *JAMA Intern. Med.* **2018**, *178*, 1544–1547. [\[CrossRef\]](#)
9. Pagano, T.P.; Loureiro, R.B.; Lisboa, F.V.; Peixoto, R.M.; Guimarães, G.A.; Cruz, G.O.; Araujo, M.M.; Santos, L.L.; Cruz, M.A.; Oliveira, E.L. Bias and unfairness in machine learning models: A systematic review on datasets, tools, fairness metrics, and identification and mitigation methods. *Big Data Cogn. Comput.* **2023**, *7*, 15. [\[CrossRef\]](#)
10. Van Giffen, B.; Herhausen, D.; Fahse, T. Overcoming the pitfalls and perils of algorithms: A classification of machine learning biases and mitigation methods. *J. Bus. Res.* **2022**, *144*, 93–106. [\[CrossRef\]](#)
11. Zhang, P.; He, Q.; Ai, X.; Ma, F. Uncovering Self-Supervised Learning: From Current Applications to Future Trends. In Proceedings of the 2023 International Conference on Power, Communication, Computing and Networking Technologies, Wuhan, China, 24–25 September 2023; pp. 1–8.
12. Rani, V.; Nabi, S.T.; Kumar, M.; Mittal, A.; Kumar, K. Self-supervised learning: A succinct review. *Arch. Comput. Methods Eng.* **2023**, *30*, 2761–2775. [\[CrossRef\]](#)
13. Zhao, Z.; Alzubaidi, L.; Zhang, J.; Duan, Y.; Gu, Y. A comparison review of transfer learning and self-supervised learning: Definitions, applications, advantages and limitations. *Expert Syst. Appl.* **2023**, *242*, 122807. [\[CrossRef\]](#)
14. Albelwi, S. Survey on self-supervised learning: Auxiliary pretext tasks and contrastive learning methods in imaging. *Entropy* **2022**, *24*, 551. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Huang, S.-C.; Pareek, A.; Jensen, M.; Lungren, M.P.; Yeung, S.; Chaudhari, A.S. Self-supervised learning for medical image classification: A systematic review and implementation guidelines. *NPJ Digit. Med.* **2023**, *6*, 74. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Purushwalkam, S.; Morgado, P.; Gupta, A. The challenges of continuous self-supervised learning. In Proceedings of the European Conference on Computer Vision, Tel Aviv, Israel, 23–27 October 2022; pp. 702–721.
17. Taherdoost, H. Blockchain and machine learning: A critical review on security. *Information* **2023**, *14*, 295. [\[CrossRef\]](#)
18. Sokolova, M.; Lapalme, G. A systematic analysis of performance measures for classification tasks. *Inf. Process. Manag.* **2009**, *45*, 427–437. [\[CrossRef\]](#)
19. Bradley, A.P. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognit.* **1997**, *30*, 1145–1159. [\[CrossRef\]](#)
20. Powers, D.M. Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. *arXiv* **2020**, arXiv:2010.16061.
21. Nahm, F.S. Receiver operating characteristic curve: Overview and practical use for clinicians. *Korean J. Anesthesiol.* **2022**, *75*, 25. [\[CrossRef\]](#) [\[PubMed\]](#)
22. de Hond, A.A.; Steyerberg, E.W.; van Calster, B. Interpreting area under the receiver operating characteristic curve. *Lancet Digit. Health* **2022**, *4*, e853–e855. [\[CrossRef\]](#)
23. Polo, T.C.F.; Miot, H.A. Use of ROC curves in clinical and experimental studies. *J. Vasc. Bras.* **2020**, *19*, e20200186. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Hajian-Tilaki, K. Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Casp. J. Intern. Med.* **2013**, *4*, 627.
25. Kwegyir-Aggrey, K.; Gerchick, M.; Mohan, M.; Horowitz, A.; Venkatasubramanian, S. The Misuse of AUC: What High Impact Risk Assessment Gets Wrong. In Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency, Chicago, IL, USA, 12–15 June 2023; pp. 1570–1583.
26. Mandrekar, J.N. Receiver operating characteristic curve in diagnostic test assessment. *J. Thorac. Oncol.* **2010**, *5*, 1315–1316. [\[CrossRef\]](#)
27. Chen, T.; Kornblith, S.; Norouzi, M.; Hinton, G. A simple framework for contrastive learning of visual representations. In Proceedings of the International Conference on Machine Learning, Virtual, 13–18 July 2020; pp. 1597–1607.

28. He, K.; Fan, H.; Wu, Y.; Xie, S.; Girshick, R. Momentum contrast for unsupervised visual representation learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 9729–9738.
29. Chiaroni, F.; Rahal, M.-C.; Hueber, N.; Dufaux, F. Self-supervised learning for autonomous vehicles perception: A conciliation between analytical and learning methods. *IEEE Signal Process. Mag.* **2020**, *38*, 31–41. [\[CrossRef\]](#)
30. Radford, A.; Kim, J.W.; Hallacy, C.; Ramesh, A.; Goh, G.; Agarwal, S.; Sastry, G.; Askell, A.; Mishkin, P.; Clark, J. Learning transferable visual models from natural language supervision. In Proceedings of the International Conference on Machine Learning, Online, 18–24 July 2021; pp. 8748–8763.
31. Kong, D.; Zhao, L.; Huang, X.; Huang, W.; Ding, J.; Yao, Y.; Xu, L.; Yang, P.; Yang, G. Self-supervised knowledge mining from unlabeled data for bearing fault diagnosis under limited annotations. *Measurement* **2023**, *220*, 113387. [\[CrossRef\]](#)
32. Shwartz-Ziv, R.; Balestrieri, R.; Kawaguchi, K.; Rudner, T.G.; LeCun, Y. An information-theoretic perspective on variance-invariance-covariance regularization. *arXiv* **2023**, arXiv:2303.00633.
33. Misra, I.; Maaten, L.v.d. Self-supervised learning of pretext-invariant representations. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 6707–6717.
34. Arora, S.; Khandeparkar, H.; Khodak, M.; Plevrakis, O.; Saunshi, N. A theoretical analysis of contrastive unsupervised representation learning. *arXiv* **2019**, arXiv:1902.09229.
35. Shurrab, S.; Duwairi, R. Self-supervised learning methods and applications in medical imaging analysis: A survey. *PeerJ Comput. Sci.* **2022**, *8*, e1045. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Liu, X.; Zhang, F.; Hou, Z.; Mian, L.; Wang, Z.; Zhang, J.; Tang, J. Self-supervised learning: Generative or contrastive. *IEEE Trans. Knowl. Data Eng.* **2021**, *35*, 857–876. [\[CrossRef\]](#)
37. Latif, S.; Rana, R.; Qadir, J.; Epps, J. Variational autoencoders for learning latent representations of speech emotion: A preliminary study. *arXiv* **2017**, arXiv:1712.08708.
38. Saxena, D.; Cao, J. Generative adversarial networks (GANs) challenges, solutions, and future directions. *ACM Comput. Surv. (CSUR)* **2021**, *54*, 1–42. [\[CrossRef\]](#)
39. Abdulrazzaq, M.M.; Ramaha, N.T.; Hameed, A.A.; Salman, M.; Yon, D.K.; Fitriyani, N.L.; Syafrudin, M.; Lee, S.W. Consequential Advancements of Self-Supervised Learning (SSL) in Deep Learning Contexts. *Mathematics* **2024**, *12*, 758. [\[CrossRef\]](#)
40. Ren, X.; Wei, W.; Xia, L.; Huang, C. A comprehensive survey on self-supervised learning for recommendation. *arXiv* **2024**, arXiv:2404.03354.
41. Ramírez, J.G.C. Advancements in Self-Supervised Learning for Remote Sensing Scene Classification: Present Innovations and Future Outlooks. *J. Artif. Intell. Gen. Sci. (JAIGS)* **2024**, *4*, 45–56.
42. Khan, M.R. Advancements in Deep Learning Architectures: A Comprehensive Review of Current Trends. *J. Artif. Intell. Gen. Sci. (JAIGS)* **2024**, *1*. [\[CrossRef\]](#)
43. Radak, M.; Lafta, H.Y.; Fallahi, H. Machine learning and deep learning techniques for breast cancer diagnosis and classification: A comprehensive review of medical imaging studies. *J. Cancer Res. Clin. Oncol.* **2023**, *149*, 10473–10491. [\[CrossRef\]](#)
44. Nielsen, M.; Wenderoth, L.; Sentker, T.; Werner, R. Self-supervision for medical image classification: State-of-the-art performance with ~ 100 labeled training samples per class. *Bioengineering* **2023**, *10*, 895. [\[CrossRef\]](#)
45. Zhai, X.; Oliver, A.; Kolesnikov, A.; Beyer, L. S4I: Self-supervised semi-supervised learning. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 1476–1485.
46. Azizi, S.; Culp, L.; Freyberg, J.; Mustafa, B.; Baur, S.; Kornblith, S.; Chen, T.; Tomasev, N.; Mitrović, J.; Strachan, P. Robust and data-efficient generalization of self-supervised machine learning for diagnostic imaging. *Nat. Biomed. Eng.* **2023**, *7*, 756–779. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Di Liello, L. Structural Self-Supervised Objectives for Transformers. *arXiv* **2023**, arXiv:2309.08272.
48. Zhou, M.; Li, Z.; Xie, P. Self-supervised regularization for text classification. *Trans. Assoc. Comput. Linguist.* **2021**, *9*, 641–656. [\[CrossRef\]](#)
49. Gururangan, S.; Marasović, A.; Swayamdipta, S.; Lo, K.; Beltagy, I.; Downey, D.; Smith, N.A. Don’t stop pretraining: Adapt language models to domains and tasks. *arXiv* **2020**, arXiv:2004.10964.
50. Sun, Y.; Wang, S.; Li, Y.; Feng, S.; Tian, H.; Wu, H.; Wang, H. Ernie 2.0: A continual pre-training framework for language understanding. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; pp. 8968–8975.
51. Karnan, H.; Yang, E.; Farkash, D.; Warnell, G.; Biswas, J.; Stone, P. STERLING: Self-Supervised Terrain Representation Learning from Unconstrained Robot Experience. In Proceedings of the 7th Annual Conference on Robot Learning, Atlanta, GA, USA, 6–9 November 2023.
52. Abbate, G.; Giusti, A.; Schmuck, V.; Celiktutan, O.; Paolillo, A. Self-supervised prediction of the intention to interact with a service robot. *Robot. Auton. Syst.* **2024**, *171*, 104568. [\[CrossRef\]](#)
53. Sermanet, P.; Lynch, C.; Chebotar, Y.; Hsu, J.; Jang, E.; Schaal, S.; Levine, S.; Brain, G. Time-contrastive networks: Self-supervised learning from video. In Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA), Brisbane, Australia, 21–25 May 2018; pp. 1134–1141.
54. Lang, C.; Braun, A.; Schillingmann, L.; Valada, A. Self-supervised multi-object tracking for autonomous driving from consistency across timescales. *IEEE Robot. Autom. Lett.* **2023**, *8*, 7711–7718. [\[CrossRef\]](#)

55. Luo, C.; Yang, X.; Yuille, A. Self-supervised pillar motion learning for autonomous driving. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 19–25 June 2021; pp. 3183–3192.
56. Bojarski, M.; Del Testa, D.; Dworakowski, D.; Firner, B.; Flepp, B.; Goyal, P.; Jackel, L.D.; Monfort, M.; Muller, U.; Zhang, J. End to end learning for self-driving cars. *arXiv* **2016**, arXiv:1604.07316.
57. Bachute, M.R.; Subhedar, J.M. Autonomous driving architectures: Insights of machine learning and deep learning algorithms. *Mach. Learn. Appl.* **2021**, *6*, 100164. [\[CrossRef\]](#)
58. Namdar, K.; Haider, M.A.; Khalvati, F. A modified AUC for training convolutional neural networks: Taking confidence into account. *Front. Artif. Intell.* **2021**, *4*, 582928. [\[CrossRef\]](#)
59. Kim, Y.; Toh, K.-A.; Teoh, A.B.J.; Eng, H.-L.; Yau, W.-Y. An online AUC formulation for binary classification. *Pattern Recognit.* **2012**, *45*, 2266–2279. [\[CrossRef\]](#)
60. Leevy, J.L.; Hancock, J.; Khoshgoftaar, T.M.; Abdollah Zadeh, A. Investigating the effectiveness of one-class and binary classification for fraud detection. *J. Big Data* **2023**, *10*, 157. [\[CrossRef\]](#)
61. Baumann, C.; Singmann, H.; Gershman, S.J.; von Helversen, B. A linear threshold model for optimal stopping behavior. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 12750–12755.
62. Djulbegovic, B.; Hozo, I.; Mayrhofer, T.; van den Ende, J.; Guyatt, G. The threshold model revisited. *J. Eval. Clin. Pract.* **2019**, *25*, 186–195. [\[CrossRef\]](#)
63. Kopsinis, Y.; Thompson, J.S.; Mulgrew, B. System-independent threshold and BER estimation in optical communications using the extended generalized gamma distribution. *Opt. Fiber Technol.* **2007**, *13*, 39–45.
64. Vanderlooy, S.; Hüllermeier, E. A critical analysis of variants of the AUC. *Mach. Learn.* **2008**, *72*, 247–262. [\[CrossRef\]](#)
65. Ferri, C.; Hernández-Orallo, J.; Flach, P.A. A coherent interpretation of AUC as a measure of aggregated classification performance. In Proceedings of the 28th International Conference on Machine Learning (ICML-11), Bellevue, WA, USA, 28 June–2 July 2011; pp. 657–664.
66. Yang, Z.; Xu, Q.; Bao, S.; He, Y.; Cao, X.; Huang, Q. Optimizing two-way partial auc with an end-to-end framework. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *45*, 10228–10246.
67. Bhat, S.; Mansoor, A.; Georgescu, B.; Panambur, A.B.; Ghesu, F.C.; Islam, S.; Packhäuser, K.; Rodríguez-Salas, D.; Grbic, S.; Maier, A. AUCReshaping: Improved sensitivity at high-specificity. *Sci. Rep.* **2023**, *13*, 21097. [\[CrossRef\]](#) [\[PubMed\]](#)
68. Sharma, R.; Ji, K.; Chen, C. AUC-CL: A Batchsize-Robust Framework for Self-Supervised Contrastive Representation Learning. In Proceedings of the The Twelfth International Conference on Learning Representations, Vienna, Austria, 7–11 May 2023.
69. Wang, S.; Li, D.; Petrick, N.; Sahiner, B.; Linguraru, M.G.; Summers, R.M. Optimizing area under the ROC curve using semi-supervised learning. *Pattern Recognit.* **2015**, *48*, 276–287. [\[CrossRef\]](#) [\[PubMed\]](#)
70. Brown, J. Classifiers and their metrics quantified. *Mol. Inform.* **2018**, *37*, 1700127. [\[CrossRef\]](#) [\[PubMed\]](#)
71. Halimu, C.; Kasem, A.; Newaz, S.S. Empirical comparison of area under ROC curve (AUC) and Mathew correlation coefficient (MCC) for evaluating machine learning algorithms on imbalanced datasets for binary classification. In Proceedings of the 3rd International Conference on Machine Learning and Soft Computing, Da Lat, Vietnam, 25–28 January 2019; pp. 1–6.
72. Ling, C.X.; Huang, J.; Zhang, H. AUC: A better measure than accuracy in comparing learning algorithms. In Proceedings of the Advances in Artificial Intelligence: 16th Conference of the Canadian Society for Computational Studies of Intelligence, AI 2003, Halifax, NS, Canada, 11–13 June 2003; Proceedings 16, 2003. pp. 329–341.
73. Yuan, Z.; Yan, Y.; Sonka, M.; Yang, T. Large-scale robust deep auc maximization: A new surrogate loss and empirical studies on medical image classification. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021; pp. 3040–3049.
74. Dong, Y.; Pachade, S.; Liang, X.; Sheth, S.A.; Giancardo, L. A self-supervised learning approach for registration agnostic imaging models with 3D brain CTA. *iScience* **2024**, *27*, 109004. [\[CrossRef\]](#) [\[PubMed\]](#)
75. Hu, R.; Chen, J.; Zhou, L. Spatiotemporal self-supervised representation learning from multi-lead ECG signals. *Biomed. Signal Process. Control* **2023**, *84*, 104772. [\[CrossRef\]](#)
76. Zhao, Y.; Li, C.; Liu, X.; Qian, R.; Song, R.; Chen, X. Patient-Specific Seizure Prediction via Adder Network and Supervised Contrastive Learning. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2022**, *30*, 1536–1547. [\[CrossRef\]](#) [\[PubMed\]](#)
77. Lu, H.; Dai, Q. A self-supervised COVID-19 CT recognition system with multiple regularizations. *Comput. Biol. Med.* **2022**, *150*, 106149. [\[CrossRef\]](#) [\[PubMed\]](#)
78. Sun, J.; Pi, P.; Tang, C.; Wang, S.-H.; Zhang, Y.-D. TSRNet: Diagnosis of COVID-19 based on self-supervised learning and hybrid ensemble model. *Comput. Biol. Med.* **2022**, *146*, 105531. [\[CrossRef\]](#)
79. Pascual, G.; Laiz, P.; García, A.; Wenzek, H.; Vitrià, J.; Seguí, S. Time-based self-supervised learning for Wireless Capsule Endoscopy. *Comput. Biol. Med.* **2022**, *146*, 105631. [\[CrossRef\]](#)
80. Wongchaisuwat, N.; Thamphithak, R.; Watunyuta, P.; Wongchaisuwat, P. Automated classification of polypoidal choroidal vasculopathy and wet age-related macular degeneration by spectral domain optical coherence tomography using self-supervised learning. *Procedia Comput. Sci.* **2023**, *220*, 1003–1008. [\[CrossRef\]](#)
81. Liu, J.; Qi, L.; Xu, Q.; Chen, J.; Cui, S.; Li, F.; Wang, Y.; Cheng, S.; Tan, W.; Zhou, Z.; et al. A Self-supervised Learning-Based Fine-Grained Classification Model for Distinguishing Malignant From Benign Subcentimeter Solid Pulmonary Nodules. *Acad. Radiol.* **2024**, in press. [\[CrossRef\]](#)

82. Xu, C.; Feng, J.; Yue, Y.; Cheng, W.; He, D.; Qi, S.; Zhang, G. A hybrid few-shot multiple-instance learning model predicting the aggressiveness of lymphoma in PET/CT images. *Comput. Methods Programs Biomed.* **2024**, *243*, 107872. [\[CrossRef\]](#)
83. Perumal, M.; Srinivas, M. DenSplitnet: Classifier-invariant neural network method to detect COVID-19 in chest CT data. *J. Vis. Commun. Image Represent.* **2023**, *97*, 103949. [\[CrossRef\]](#)
84. Manna, S.; Bhattacharya, S.; Pal, U. Self-supervised representation learning for detection of ACL tear injury in knee MR videos. *Pattern Recognit. Lett.* **2022**, *154*, 37–43. [\[CrossRef\]](#)
85. Xu, R.; Hao, R.; Huang, B. Efficient surface defect detection using self-supervised learning strategy and segmentation network. *Adv. Eng. Inform.* **2022**, *52*, 101566. [\[CrossRef\]](#)
86. Zhou, S.; Tian, S.; Yu, L.; Wu, W.; Zhang, D.; Peng, Z.; Zhou, Z. Growth threshold for pseudo labeling and pseudo label dropout for semi-supervised medical image classification. *Eng. Appl. Artif. Intell.* **2024**, *130*, 107777. [\[CrossRef\]](#)
87. Zhou, S.; Tian, S.; Yu, L.; Wu, W.; Zhang, D.; Peng, Z.; Zhou, Z.; Wang, J. FixMatch-LS: Semi-supervised skin lesion classification with label smoothing. *Biomed. Signal Process. Control* **2023**, *84*, 104709. [\[CrossRef\]](#)
88. Uegami, W.; Bychkov, A.; Ozasa, M.; Uehara, K.; Kataoka, K.; Johkoh, T.; Kondoh, Y.; Sakanashi, H.; Fukuoka, J. MIXTURE of human expertise and deep learning—Developing an explainable model for predicting pathological diagnosis and survival in patients with interstitial lung disease. *Mod. Pathol.* **2022**, *35*, 1083–1091. [\[CrossRef\]](#)
89. Zhao, B.; Deng, W.; Li, Z.H.; Zhou, C.; Gao, Z.; Wang, G.; Li, X. LESS: Label-efficient multi-scale learning for cytological whole slide image screening. *Med. Image Anal.* **2024**, *94*, 103109. [\[CrossRef\]](#)
90. Orlandic, L.; Teijeiro, T.; Atienza, D. A semi-supervised algorithm for improving the consistency of crowdsourced datasets: The COVID-19 case study on respiratory disorder classification. *Comput. Methods Programs Biomed.* **2023**, *241*, 107743. [\[CrossRef\]](#)
91. Chakravarty, A.; Emre, T.; Leingang, O.; Riedl, S.; Mai, J.; Scholl, H.P.N.; Sivaprasad, S.; Rueckert, D.; Lotery, A.; Schmidt-Erfurth, U.; et al. Morph-SSL: Self-Supervision with Longitudinal Morphing for Forecasting AMD Progression from OCT Volumes. *IEEE Trans. Med. Imaging* **2024**. [\[CrossRef\]](#)
92. Zhang, P.; Hu, X.; Li, G.; Deng, L. AntiViralDL: Computational Antiviral Drug Repurposing Using Graph Neural Network and Self-Supervised Learning. *IEEE J. Biomed. Health Inform.* **2024**, *28*, 548–556. [\[CrossRef\]](#)
93. Li, T.; Guo, Y.; Zhao, Z.; Chen, M.; Lin, Q.; Hu, X.; Yao, Z.; Hu, B. Automated Diagnosis of Major Depressive Disorder With Multi-Modal MRIs Based on Contrastive Learning: A Few-Shot Study. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2024**, *32*, 1566–1576. [\[CrossRef\]](#)
94. Huang, C.; Xu, Q.; Wang, Y.; Wang, Y.; Zhang, Y. Self-Supervised Masking for Unsupervised Anomaly Detection and Localization. *IEEE Trans. Multimed.* **2023**, *25*, 4426–4438. [\[CrossRef\]](#)
95. Wang, Q.; Chen, K.; Dou, W.; Ma, Y. Cross-Attention Based Multi-Resolution Feature Fusion Model for Self-Supervised Cervical OCT Image Classification. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2023**, *20*, 2541–2554. [\[CrossRef\]](#)
96. Yang, Z.; Liang, J.; Xu, Y.; Zhang, X.Y.; He, R. Masked Relation Learning for DeepFake Detection. *IEEE Trans. Inf. Forensics Secur.* **2023**, *18*, 1696–1708. [\[CrossRef\]](#)
97. Zhu, H.; Wang, J.; Zhao, Y.P.; Lu, M.; Shi, J. Contrastive Multi-View Composite Graph Convolutional Networks Based on Contribution Learning for Autism Spectrum Disorder Classification. *IEEE Trans. Biomed. Eng.* **2023**, *70*, 1943–1954. [\[CrossRef\]](#)
98. Yu, J.G.; Wu, Z.; Ming, Y.; Deng, S.; Wu, Q.; Xiong, Z.; Yu, T.; Xia, G.S.; Jiang, Q.; Li, Y. Bayesian Collaborative Learning for Whole-Slide Image Classification. *IEEE Trans. Med. Imaging* **2023**, *42*, 1809–1821. [\[CrossRef\]](#)
99. Kragh, M.F.; Rimestad, J.; Lassen, J.T.; Berntsen, J.; Karstoft, H. Predicting Embryo Viability Based on Self-Supervised Alignment of Time-Lapse Videos. *IEEE Trans. Med. Imaging* **2022**, *41*, 465–475. [\[CrossRef\]](#)
100. Luo, J.; Lin, J.; Yang, Z.; Liu, H. SMD Anomaly Detection: A Self-Supervised Texture–Structure Anomaly Detection Framework. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 1–11. [\[CrossRef\]](#)
101. Huang, H.; Wu, R.; Li, Y.; Peng, C. Self-Supervised Transfer Learning Based on Domain Adaptation for Benign-Malignant Lung Nodule Classification on Thoracic CT. *IEEE J. Biomed. Health Inform.* **2022**, *26*, 3860–3871. [\[CrossRef\]](#)
102. Schmidt, A.; Silva-Rodríguez, J.; Molina, R.; Naranjo, V. Efficient Cancer Classification by Coupling Semi Supervised and Multiple Instance Learning. *IEEE Access* **2022**, *10*, 9763–9773. [\[CrossRef\]](#)
103. Kim, K.S.; Oh, S.J.; Cho, H.B.; Chung, M.J. One-Class Classifier for Chest X-Ray Anomaly Detection via Contrastive Patch-Based Percentile. *IEEE Access* **2021**, *9*, 168496–168510. [\[CrossRef\]](#)
104. Tardy, M.; Mateus, D. Looking for Abnormalities in Mammograms With Self- and Weakly Supervised Reconstruction. *IEEE Trans. Med. Imaging* **2021**, *40*, 2711–2722. [\[CrossRef\]](#)
105. Godson, L.; Alemi, N.; Nsengimana, J.; Cook, G.P.; Clarke, E.L.; Treanor, D.; Bishop, D.T.; Newton-Bishop, J.; Gooya, A.; Magee, D. Immune subtyping of melanoma whole slide images using multiple instance learning. *Med. Image Anal.* **2024**, *93*, 103097. [\[CrossRef\]](#)
106. Bai, Y.; Li, W.; An, J.; Xia, L.; Chen, H.; Zhao, G.; Gao, Z. Masked autoencoders with handcrafted feature predictions: Transformer for weakly supervised esophageal cancer classification. *Comput. Methods Programs Biomed.* **2024**, *244*, 107936. [\[CrossRef\]](#)

107. Ali, M.K.; Amin, B.; Maud, A.R.; Bhatti, F.A.; Sukhia, K.N.; Khurshid, K. Hyperspectral target detection using self-supervised background learning. *Adv. Space Res.* **2024**, *74*, 628–646. [[CrossRef](#)]
108. Bastos, M. Human-Centered Design of a Semantic Annotation Tool for Breast Cancer Diagnosis. Available online: https://www.researchgate.net/profile/Francisco-Maria-Calisto/publication/379311291_Human-Centered_Design_of_a_Semantic_Annotation_Tool_for_Breast_Cancer_Diagnosis/links/66041361390c214cfd14da37/Human-Centered-Design-of-a-Semantic-Annotation-Tool-for-Breast-Cancer-Diagnosis.pdf (accessed on 13 July 2024).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.